

# Maximal averages along a planar vector field depending on one variable \*

Michael Bateman <sup>†</sup>

## Abstract

We prove (essentially) sharp  $L^2$  estimates for a restricted maximal operator associated to a planar vector field that depends only on the horizontal variable. The proof combines an understanding of such vector fields from earlier work of the author with a result of Nets Katz on directional maximal operators.

## 1 Introduction

We prove an estimate on the  $L^2$  norm of a certain maximal operator related to vector fields depending on only one variable. The author has previously established bounds on the  $L^p$  norm of this operator; interpolating these with the  $L^2$  bounds in this paper yields (essentially) sharp  $L^p$  estimates. This theorem is loosely related to the problem of bounding Hilbert transforms along a vector field. Defining the maximal operator requires a bit of notation, which we present below.

### 1.1 Averages over rectangles

We start by defining a maximal operator for any collection of rectangles  $\mathcal{R}$ :

$$M_{\mathcal{R}}f(x) = \sup_{x \in R \in \mathcal{R}} \frac{1}{|R|} \int_R f.$$

Let  $v: \mathbb{R}^2 \rightarrow [0, 1]$ . For any rectangle  $R$ , let  $L(R)$  denote the length of  $R$ ,  $w(R)$  the width of  $R$ , and let  $\theta(R)$  be the interval of width  $\frac{w(R)}{L(R)}$  centered at the slope of the long side of  $R$ . Now let  $V_R = \{p \in R: v(p) \in \theta(R)\}$ . Next we define the collection of rectangles concerning us. Fix two parameters  $0 < w \leq 1$  and  $0 < \delta \leq 1$ , and define

$$\mathcal{R}_\delta = \{\text{rectangles } R \text{ of width } w: |V(R)| \geq \delta|R|\}.$$

---

\*This work was supported by NSF grant DMS-0902490. 2010 Mathematics Subject Classification: Primary: 42B25, Secondary: 42B20 .

<sup>†</sup>Department of Mathematics, UCLA, Box 951555, Los Angeles, CA 90095-1555  
([bateman@math.ucla.edu](mailto:bateman@math.ucla.edu).)

Note that the definition of  $\mathcal{R}_\delta$  depends on the vector field  $v$ ; we will suppress this dependence.

**Theorem 1.** *Suppose  $v: \mathbb{R}^2 \rightarrow [0, 1]$  depends on one variable, i.e.,  $v(a, b) = v(a)$ . Then for any  $f \in L^2(\mathbb{R}^2)$ ,*

$$\|M_{\mathcal{R}_\delta} f\|_2 \lesssim \left( \log \frac{1}{\delta} \right)^{\frac{3}{2}} \|f\|_2.$$

This estimate can be interpolated with the obvious  $L^\infty$  bound on  $M_{\mathcal{R}_\delta}$  to obtain logarithmic bounds when  $p \geq 2$ . Additionally, we have the following corollary:

**Corollary 2.** *Under the same hypotheses as the theorem, when  $p \in (1, 2)$  we have for  $f \in L^p(\mathbb{R}^2)$ ,*

$$\|M_{\mathcal{R}_\delta} f\|_p \lesssim \left( \log \frac{1}{\delta} \right)^{3(1-\frac{1}{p})} \frac{1}{\delta^{\frac{2}{p}-1}} \|f\|_p.$$

We remark that the theorem here is for rectangles of a fixed width  $w$ . It is not clear that the argument here generalizes to the situation of rectangles with arbitrary width. However, a theorem in that setting may play a role in the study of Hilbert transforms along a one-variable vector field. Motivation for studying maximal averages comes from differentiation theory; this connection has been known for some time. More recently, maximal theorems of this flavor (with the density parameter  $\delta$ ) have been connected to the study of Hilbert transforms along a vector field. See [5], [6] for more on this connection.

We prove the theorem by combining ideas developed by Nets Katz in the study of directional maximal operators ([3], [4]) together with the understanding of one-variable vector fields obtained by the author in [1]. It is likely that the exponent on the logarithm is not sharp, but we do not pursue that idea here. (For example, the argument to obtain estimate (3) below is rather crude.) We note however that the operator norm is at least  $\sqrt{\log \frac{1}{\delta}}$ . This can be seen by considering the slope field  $v(x, y) = x$  defined on  $[0, 1]^2$ . Then for  $w = \delta$ ,  $\mathcal{R}_\delta$  contains (at least) all rectangles of length 1 projecting vertically to  $[0, 1]$  with slope in  $[0, 1]$ . Now we can construct Kakeya-type sets using rectangles from  $\mathcal{R}_\delta$ . Letting  $f$  be the characteristic function of such a set shows that  $\|M_{\mathcal{R}_\delta}\|_2 \gtrsim \sqrt{\log \frac{1}{\delta}}$ . Similarly, the  $L^p$  estimates given in the corollary are sharp up to logarithmic factors; this can be seen by again considering  $v(x, y) = x$  and letting  $f$  be the characteristic function of a  $\delta \times \delta$  square. In this setup  $M_{\mathcal{R}_\delta} \gtrsim \delta$  on a set of measure approximately one.

## 2 Outline of proof

Recall that all rectangles in question have a fixed width  $w$ . By a standard reduction (see [1]), we may assume that the slope of each rectangle of length

$2^k w$  is in the discrete set

$$S_k = \left\{ \frac{j + \frac{1}{2}}{2^k} : j \in \{0, 1, \dots, 2^k - 1\} \right\}.$$

Further, we may assume our “rectangles” are actually parallelograms projecting to dyadic intervals. It will be convenient later to assume all rectangles in  $\mathcal{R}_\delta$  live in a bounded region, which we take to be the unit square. We may do so by (say) approximating with finite subcollections of  $\mathcal{R}_\delta$ . Next we linearize the maximal operator. That is, for each  $x \in \mathbb{R}^2$  we choose a rectangle  $R \in \mathcal{R}_\delta$  that nearly achieves the supremum in the definition of the maximal operator. We will call this rectangle  $\rho(x)$ . It is possible that there is no  $R \in \mathcal{R}_\delta$  containing  $x$ ; let  $X$  denote the set of points with this property. This gives us a map  $\rho: \mathbb{R}^2 \setminus X \rightarrow \mathcal{R}_\delta$  and a linear operator defined by

$$T_\rho f(x) = \frac{1}{|\rho(x)|} \int_{\rho(x)} f$$

for  $x \notin X$ , and  $T_\rho f(x) = 0$  for  $x \in X$ . Without loss of generality, we will take this region to be the unit square. To prove Theorem 1, it suffices to prove the same bounds on the linear operators  $T_\rho$  independent of the choice function  $\rho$ . To do this, we decompose the operator in a certain way depending on the vector field and the linearization  $\rho$ , and apply the Cotlar-Stein lemma. From now on, we consider the function  $\rho$  to be fixed, and we write  $T = T_\rho$ . We recall a variant of the Cotlar-Stein lemma.

**Lemma 3.** [Cotlar-Stein] Suppose  $\{T_j\}_{j=-\infty}^\infty$  is a sequence of linear operators acting on a Hilbert space  $\mathbf{H}$ , and let  $T = \sum_{j=-\infty}^\infty T_j$ . Assume that  $a: \mathbb{Z} \rightarrow \mathbb{R}$  is such that for all  $j, k$ ,

$$\|T_j T_k^*\| \leq a(j - k) \tag{1}$$

$$\|T_j^* T_k\| = 0.$$

Then

$$\|T\| \leq a(0)^{\frac{1}{2}} \left( \sum \sqrt{a(j)} \right)^{\frac{1}{2}}.$$

A straightforward modification of the proof in [2] gives the result claimed here. The following lemma shows how we will satisfy the hypotheses of the Cotlar-Stein lemma.

**Lemma 4.** Suppose  $v$  is a vector field depending on one variable. There exist pairwise disjoint sets  $A_1, A_2, A_3, \dots$ , such that if we define

$$T_j f(x) = \mathbf{1}_{A_j}(x) T f(x)$$

for  $j = 1, 2, 3, \dots$ , then for all  $j, k$ , we have

$$\|T_j T_k^*\| \lesssim \log \left( \frac{1}{\delta} \right)^2 \tag{2}$$

$$\|T_j T_k^*\| \lesssim \frac{1}{\delta} 2^{-|j-k|}, \quad (3)$$

and

$$\|T_j^* T_k\| = 0.$$

We will use estimate (2) when  $|j - k| \lesssim \log \frac{1}{\delta}$  and estimate (3) otherwise. We remark that  $\|T_j^* T_k\| = 0$  automatically for  $j \neq k$ , since in this case  $A_j \cap A_k = \emptyset$ . Also,  $\|T_j^* T_j\| = \|T_j T_j^*\|$ . With this in mind, Theorem 1 follows easily from the previous two lemmas by letting

$$a(n) = C(\log \frac{1}{\delta})^2 \text{ for } n \leq C \log \frac{1}{\delta},$$

$$a(C \log \frac{1}{\delta} + n) = 2^{-n} \text{ for } n \geq 0,$$

and

$$a(n) = a(-n) \text{ for } n < 0.$$

Applying the Cotlar-Stein lemma, we see that

$$\|T\| \leq a(0)^{\frac{1}{2}} \left( \sum \sqrt{a(n)} \right)^{\frac{1}{2}} \lesssim (\log \frac{1}{\delta})^{\frac{3}{2}}.$$

## 2.1 Agenda

The remainder of the paper is devoted to proving Lemma 4. In Section 3, we present the large components of the proof of Lemma 4 and show how they imply (2). In Section 5, we show how Lemma 7 in Section 3 implies (3). In Sections 4 and 6, we prove the lemmas from Section 3.

## 2.2 Notation

If  $A$  is a set, we write  $\mathbf{1}_A$  to denote the characteristic function of  $A$ . We write  $C$  to denote universal constants that may vary from one appearance to the next. We write  $x \lesssim y$  to mean  $x \leq Cy$ . If  $\mathcal{C}$  is a collection of sets, we write

$$\mathbf{sh}(\mathcal{C}) = \bigcup_{C \in \mathcal{C}} C.$$

**sh** stands for ‘‘shadow’’. If  $A \subseteq \mathbb{R}^2$ , we write  $\pi_1(A)$  to denote the projection of  $A$  onto the horizontal axis, and  $\pi_2(A)$  to denote the projection of  $A$  onto the vertical axis.

### 3 The main ingredients

In this section we present the statements of the two most substantial ingredients needed for the proof of Lemma 4. The first concerns a prototype of the operator  $M_{\mathcal{R}_\delta}$  defined above, which is closely related to the maximal operator over  $\sim \frac{1}{\delta}$  arbitrary directions. A more precise definition is given below. The second key lemma is the inductive step in a stopping time argument. It tells us how to define the sets  $A_j$  needed for the decomposition of our operator  $T$  by identifying the intervals on which rectangles of many different directions might be chosen.

#### 3.1 Statements

We start with a definition.

**Definition 5.** *We say a collection of rectangles  $\mathcal{R}$  is good if whenever  $R_1, R_2 \in \mathcal{R}$  are such that  $\pi_1(R_1) = \pi_1(R_2)$ , the slope of  $R_1$  equals the slope of  $R_2$ ; and if  $M_{\mathcal{R}}$  is weak (1,1).*

**Theorem 6.** *Let  $v$  be a vector field, and let  $N \geq 2$  be an integer. Suppose  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_N$  are good collections of rectangles. Let  $\mathcal{R} = \bigcup_{j=1}^N \mathcal{R}_j$ . Then*

$$\|M_{\mathcal{R}}f\|_2 \lesssim \log N \|f\|_2.$$

This theorem is essentially due to Katz, in [3]. The only difference between this theorem and his is that we allow ourselves to average over  $N$  “good” collections rather than  $N$  different directions. The theorem stated here does not actually follow from Katz’s statement, but rather his proof. We include the proof at the end. The key point is that if  $\mathcal{R}_j$  is a good collection of rectangles with fixed width, then  $M_{\mathcal{R}_j}$  is weak-type (1,1). The other part of goodness is more of a convenience.

**Lemma 7.** *Let  $I \subseteq [0, 1]$  be a dyadic interval. Let  $E \subseteq I \times [0, 1]$  be such that if  $x \in E$ , then  $\pi_1(\rho(x)) \subseteq I$ . There exist collections  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\frac{3}{\delta}}$ , a collection of disjoint intervals  $\mathcal{I}_I$ , and sets  $E_{good}$  and  $E_{bad}$  such that*

1.  $\mathbf{sh}(\mathcal{I}_I) \subseteq I$ ,
2.  $|\mathbf{sh}(\mathcal{I}_I)| \leq \frac{1}{2}|I|$ ,
3. each  $\mathcal{R}_i$  is a good collection,
4.  $E_{good} \cap E_{bad} = \emptyset$ ,
5.  $E_{good} \cup E_{bad} = E$ ,
6. for  $x \in E_{good}$ , we have  $\rho(x) \in \mathcal{R}_i$  for some  $i \in \{1, 2, \dots, \frac{3}{\delta}\}$ ,
7. and for  $x \in E_{bad}$ , we have  $\pi_1(\rho(x)) \subseteq \mathbf{sh}(\mathcal{I}_I)$ .

### 3.2 Defining the sets $A_j$ used in the decomposition of $T$

We now use Lemma 7 to construct the sets  $A_j$  mentioned in Lemma 4. The last point in the lemma above guarantees that if  $x \in E_{bad}$ , then  $x \in \mathbf{sh}(\mathcal{I}_I) \times [0, 1]$ . Let  $\mathcal{I}_0 = \{[0, 1]\}$  and let  $E_0 = [0, 1]^2$ . Now suppose we have constructed the collections  $\mathcal{I}_0, \mathcal{I}_1, \dots, \mathcal{I}_j$  and the sets  $E_0, E_1, \dots, E_j$ . For each  $I \in \mathcal{I}_j$ , define  $E_{j,I} = E_j \cap (I \times [0, 1])$ . Apply the previous lemma to the intervals  $I \subseteq \mathcal{I}_j$  with the sets  $E_{j,I}$ . Define

$$\mathcal{I}_{j+1} = \bigcup_{I \in \mathcal{I}_j} \mathcal{I}_I,$$

and

$$E_{j+1} = \bigcup_{I \in \mathcal{I}_j} (E_{j,I})_{bad},$$

and

$$A_j = \bigcup_{I \in \mathcal{I}_j} (E_{j,I})_{good}.$$

These are the sets  $A_j$  used in the decomposition of our operator  $T$ . By construction, we have for each  $I \in \mathcal{I}_j$  and  $k \geq j$ ,

$$|I \cap \mathbf{sh}(\mathcal{I}_k)| \leq 2^{-|j-k|} |I|.$$

This is the key fact needed to prove the estimate (3) used for  $j, k$  far apart. The proof of (3) occupies Section 5.

### 3.3 Proof of estimate (2)

Note that for each fixed  $j$ , there exist good collections  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_{\frac{3}{\delta}}$  such that  $\rho(x) \in \bigcup_{i=1}^{\frac{3}{\delta}} \mathcal{R}_i$  for all  $x \in A_j$ . This fact, together with Theorem 6, is already enough to establish the estimate (2): Each  $T_j$  is controlled by a maximal operator  $M_{\mathcal{R}}$  where

$$\mathcal{R} = \bigcup_{i=1}^{\frac{3}{\delta}} \mathcal{R}_i$$

and each  $\mathcal{R}_i$  is a good collection. Hence we may apply Lemma 6 to obtain the estimate

$$\|T_j T_k^*\|_2 \leq \|T_j\|_2 \|T_k\|_2 \lesssim \left( \log \frac{1}{\delta} \right)^2.$$

## 4 Proof of the iterative Lemma 7

We begin by introducing some notation that will help us describe the collection  $\mathcal{I}_I$  in the statement of the lemma. Recall that all rectangles in question have a fixed width  $w$ . Also recall that the slope of each rectangle is in the discrete set defined at the beginning of Section 2 and that our “rectangles” are actually parallelograms projecting to dyadic intervals. For each dyadic interval  $J \subseteq I$  and any  $s \in S$ , we define

$$G_{J,s} = \{a \in J : v(a) \in [s - \frac{w}{|J|}, s + \frac{w}{|J|})\}$$

$$S(J) = \{s \in S : |G_{J,s}| \geq \delta|J|\}.$$

In the rest of this section we will abuse notation and write  $s$  to denote the dyadic interval centered at  $s$ . The convenience of this will be apparent throughout the section.  $S(J)$  is the set of allowable slope for rectangles projecting vertically to  $J$ . For  $J$  dyadic with  $J \subseteq I$ , we will define a set of slopes  $T(J)$  as follows. The definition is inductive, starting with the largest interval and then moving to its subintervals. First for  $I$ , the largest interval, define

$$T(I) = S(I).$$

Note that  $T(I)$  is just the set of allowable slopes for the interval  $I$ . (Recall that the allowable slopes for an interval are those that are at least  $\delta$ -popular.) Now for smaller intervals  $J$ , we will define  $T(J)$  similarly, except that we will not include slopes that have been used by an ancestor of  $J$ . (By “ancestor”, we mean another dyadic interval containing  $J$ .) More precisely, having defined  $T(K)$  for  $K \supsetneq J$ , define

$$T(J) = \{s \in S(J) : s \not\supseteq s' \text{ for any } s' \in T(K), K \supsetneq J\}.$$

For  $s \in T(J)$ , let

$$\mu_J^s = |G_{J,s}|;$$

otherwise, let  $\mu_J^s = 0$ ; and let

$$\mu_J = \sum_{s \in T(J)} \mu_J^s.$$

It is straightforward to check that

$$\sum_{J \subseteq I} \mu_J \leq |I| \tag{4}$$

since  $G_{J,s} \cap G_{J',s'} = \emptyset$  for  $s \in T(J)$  and  $s' \in T(J')$  unless  $J = J'$  and  $s = s'$ .

We now define the collection  $\mathcal{I}_I$  mentioned in the statement of the lemma. Let  $\mathcal{I}_I$  be the collection of maximal subintervals  $I'$  of  $I$  for which

$$\sum_{I' \subseteq K \subseteq I} \frac{\mu_K}{|K|} \geq 2. \quad (5)$$

We remark that

$$|\bigcup \mathcal{I}_I| \leq |\{a \in I: \sum_{K \subseteq I} \frac{\mu_K}{|K|} \mathbf{1}_K(a) \geq 2\}| \quad (6)$$

$$\leq \frac{1}{2} |I| \quad (7)$$

by Chebyshev's inequality and the Carleson condition (4). This proves the second claim of the lemma. Of course the first claim is true by construction. Let

$$\Theta = \{(J, s): J \subseteq I \text{ and } s \in T(J)\},$$

$$\Theta_{bad} = \{(J, s) \in \Theta: J \subseteq I' \text{ for some } I' \in \mathcal{I}_I\}$$

and let

$$\Theta_{good} = \Theta \setminus \Theta_{bad}.$$

The following partial order on pairs in  $\Theta$  will be useful: we write

$$(J, s) \leq (J', s')$$

whenever either  $J = J'$  and the center of  $s$  is less than or equal to the center of  $s'$ , or  $J \subsetneq J'$ . Note that if  $J \cap J' \neq \emptyset$ , then  $(J, s)$  and  $(J', s')$  are comparable under the relation  $\leq$ . Of course we will write  $(J, s) < (J', s')$  to mean  $(J, s) \leq (J', s')$  but  $(J, s) \neq (J', s')$ . Define the *children* of a pair  $(J', s')$  to be all pairs  $(J, s) < (J', s')$  that are maximal with respect to this property. (I.e., there is no pair  $(J'', s'')$  such that  $(J, s) < (J'', s'') < (J', s')$ .) Let  $C(J, s)$  denote the set of children of  $(J, s)$ . Now we sort elements of  $\Theta_{good}$  inductively. Define  $\Omega_0$  to be the set of maximal elements of  $\Theta_{good}$ . Now having defined  $\Omega_0, \Omega_1, \dots, \Omega_n$ , define

$$\Omega_{n+1} = \Theta_{good} \cap \left( \bigcup_{(J, s) \in \Omega_n} C(J, s) \right).$$

Now we let

$$F_n = \{x \in E: \exists (J, s) \in \Omega_n \text{ with } \pi_1(\rho(x)) \subseteq J \text{ and } \text{slope}(\rho(x)) \supseteq s\},$$

$$\mathcal{R}_n = \{\rho(x): x \in F_n\},$$

and define

$$E_{good} = \bigcup_{n=1}^{\infty} F_n,$$

$$E_{bad} = E \setminus E_{good}.$$

This proves claims 4,5, and 6 of the lemma by construction. Note that if  $slope(\rho(x)) \in S(\pi_1(\rho(x)))$ , then there exists  $(J, s) \in \Theta$  such that  $\pi_1(\rho(x)) \subseteq J$  and  $slope(\rho(x)) \supseteq s$ . If  $x \in E_{bad}$ , then this  $(J, s) \notin \Theta_{good}$ . Hence  $(J, s) \in \Theta_{bad}$ , so  $\pi_1(x) \in \pi_1(\rho(x)) \subseteq \mathbf{sh}(\mathcal{I}_I)$ . This proves 7. To complete the proof of the lemma, it is enough to establish the following two claims:

**Claim 8.**  $\Omega_{\frac{3}{8}}$  is empty. (From this it follows that  $F_{\frac{3}{8}}$  is empty.)

**Claim 9.** For each  $n$ ,  $\mathcal{R}_n$  is a good collection.

Recall that *good* collections are defined in Definition 5.

*Proof of Claim 9.* First note that if  $slope(\rho(x)) \in S(\pi_1(\rho(x)))$ , then there exists  $(J, s) \in \Theta$  such that  $\pi_1(\rho(x)) \subseteq J$  and  $slope(\rho(x)) \supseteq s$ . Now note that if  $\mathcal{R}_0$  is a collection of rectangles such that  $slope(R) \supseteq s$  for all  $R \in \mathcal{R}_0$ , then  $M_{\mathcal{R}_0}$  is weak (1,1). This is because all rectangles in  $\mathcal{R}_0$  essentially point in the same direction. Similarly, if we have a disjoint collection of intervals  $\mathcal{J}$  and slopes  $\{s_J\}_{J \in \mathcal{J}}$  such that for each  $R \in \mathcal{R}_0$ , we have  $J \in \mathcal{J}$  such that  $\pi_1(R) \subseteq J$  and  $slope(R) \supseteq s_J$ , then  $M_{\mathcal{R}_0}$  is again weak (1,1).

Hence the claim follows immediately from the following fact: For every  $n = 0, 1, 2, \dots$ , if  $(J_1, s_1), (J_2, s_2) \in \Omega_n$  with  $(J_1, s_1) \neq (J_2, s_2)$ , then  $J_1 \cap J_2 = \emptyset$ .

This fact follows from an easy induction argument: Since  $\Omega_0$  contains only maximal elements in  $\Theta_{good}$ , we cannot have any distinct  $(J_1, s_1), (J_2, s_2) \in \Omega_0$  with  $J_1 \subseteq J_2$ , because in that case it is not possible for both  $(J_1, s_1)$  and  $(J_2, s_2)$  to be maximal. Now suppose the claim is true for distinct pairs in  $\Omega_{n-1}$ , and suppose  $(J_1, s_1), (J_2, s_2) \in \Omega_n$ . By definition of  $\Omega_n$ , there exist  $(J'_1, s'_1), (J'_2, s'_2) \in \Omega_{n-1}$  such that  $(J_i, s_i) \leq (J'_i, s'_i)$  for  $i = 1, 2$ . This implies, in particular, that  $J_i \subseteq J'_i$  for  $i = 1, 2$ . By our induction hypothesis, we know that either  $J'_1$  does not intersect  $J'_2$ , or  $(J'_1, s'_1) = (J'_2, s'_2)$ . In the first case, it is obvious that  $J_1 \cap J_2 = \emptyset$ . In the second case, we argue as we did in the  $n = 0$  case: if, say,  $J_1 \subseteq J_2$ , then it is not possible for both of  $(J_1, s_1)$  and  $(J_2, s_2)$  to be maximal children of  $(J'_1, s'_1)$ . This proves the claim.  $\square$

*Proof of Claim 8.* We begin by defining, for any dyadic  $K \subseteq I$ ,

$$\Theta_K = \{(J, s) : K \subseteq J \subseteq I \text{ and } s \in T(J)\}.$$

Note that if  $s \in T(J)$ , then  $\frac{\mu_J^s}{|J|} \geq \delta$ , so

$$\begin{aligned}\#\Theta_K) &= \sum_{K \subseteq J \subseteq I} \#(T(J)) \\ &= \sum_{K \subseteq J \subseteq I} \sum_{s \in T(J)} 1 \\ &\leq \frac{1}{\delta} \sum_{K \subseteq J \subseteq I} \sum_{s \in T(J)} \frac{\mu_J^s}{|J|} \\ &= \frac{1}{\delta} \sum_{K \subseteq J \subseteq I} \frac{\mu_J}{|J|}.\end{aligned}$$

If the claim were false, then there would be a sequence

$$(J_1, s_1) > (J_2, s_2) > \cdots > (J_{\frac{3}{\delta}}, s_{\frac{3}{\delta}})$$

with  $(J_i, s_i) \in \Theta_{good}$  for  $i = 1, 2, \dots, \frac{3}{\delta}$ . But this implies

$$\frac{3}{\delta} \leq \#\left(\Theta_{J_{\frac{3}{\delta}}}\right) \leq \frac{1}{\delta} \sum_{J_{\frac{3}{\delta}} \subseteq K \subseteq I} \frac{\mu_K}{|K|},$$

which is impossible since  $J_{\frac{3}{\delta}} \not\subseteq I'$  for any  $I' \in \mathcal{I}_I$ . See the definition of  $\mathcal{I}_I$  in (5). This proves the claim.  $\square$

## 5 Proof of estimate (3)

In this section, we establish the estimate (3). Recall that we use this estimate when  $|j - k|$  is rather large. To prove it, we take advantage of the rapid decay of  $|J \cap \mathbf{sh}(\mathcal{I}_k)|$  whenever  $J \in \mathcal{I}_j$  and  $k$  is much larger than  $j$ . Because of this decay, we have that rectangles chosen by points in  $A_k$  will only be able to intersect rectangles  $R$  chosen by points in  $A_j$  on very small subsets of  $R$ . Essentially all of the analysis of this section takes place on a fixed interval  $J \in \mathcal{I}_j$ . We formalize these ideas below. In this section, we use notation from Sections 3 and 4. The reader may wish to ignore the dependence on  $J$  in some of the notation below and imagine that  $\mathcal{I}_j$  consists of a single interval.

Note that  $T_j T_k^* = (T_k T_j^*)^*$ , so it is enough to control  $\|T_k T_j^*\|$  in the case  $j \leq k$ . So fix  $j$  and  $k$  with  $j \leq k$ . Recall that

$$T_j f(x) = \mathbf{1}_{A_j}(x) \frac{1}{\rho(x)} \int_{\rho(x)} f,$$

where  $\rho$  is a fixed linearizing function. Fix  $J \in \mathcal{I}_j$ . By Lemma 7 and the definition of  $A_j$  following the lemma, we know there are collections  $\mathcal{R}_1^j, \mathcal{R}_2^j, \dots, \mathcal{R}_{\frac{3}{\delta}}^j$  such that if  $x \in A_j$ , then  $\rho(x) \in \cup_n \mathcal{R}_n^j$ . Further,  $\pi(\rho(x)) \subseteq \mathbf{sh}(\mathcal{I}_j)$  for all  $x \in A_j$ . Let

$$A_{j,J} = \{x \in A_j \cap (J \times [0, 1])\}$$

and for each  $n = 1, 2, \dots, \frac{3}{\delta}$ , let

$$A_{j,J,n} = \{x \in A_{j,J} : \rho(x) \in \mathcal{R}_n^j\}.$$

With this notation, if  $x \in A_{j,J}$  we define

$$T_{j,J,n}f(x) = \mathbf{1}_{A_{j,J,n}}(x)T_j f(x)$$

and

$$T_{j,J}f(x) = \sum_{n=1}^{\frac{3}{\delta}} T_{j,J,n}f(x).$$

Note that

$$T_j = \sum_{J \in \mathcal{I}_j} T_{j,J},$$

and that  $T_{j,J}^*f = T_{j,J}^*(f\mathbf{1}_{J \times [0,1]})$  is supported on  $J \times [0, 1]$ . To prove the estimate (3), it is enough to prove

$$\|T_k T_{j,J,n}^*f\| \lesssim 2^{-|j-k|} \quad (8)$$

for every  $J \in \mathcal{I}_j$  and every  $n \in \{1, 2, \dots, \frac{3}{\delta}\}$ , because then

$$\begin{aligned} \|T_k T_{j,J}^*f\|_2 &\leq \sum_{J \in \mathcal{I}_j} \sum_{n=1}^{\frac{3}{\delta}} \|T_k T_{j,J,n}^*(f\mathbf{1}_{J \times [0,1]})\|_2 \\ &\leq 2^{-|j-k|} \sum_{J \in \mathcal{I}_j} \sum_{n=1}^{\frac{3}{\delta}} \|(f\mathbf{1}_{J \times [0,1]})\|_2 \\ &\lesssim 2^{-|j-k|} \frac{1}{\delta} \|f\|_2. \end{aligned}$$

To prove (8), and hence (3), it is enough to prove the following two claims:

**Claim 10.** *For each  $n = 1, 2, \dots, \frac{3}{\delta}$ , and each  $x \in A_k$ ,*

$$T_k T_{j,J,n}^*f(x) \leq M_2 T_{j,J,n}^*f(x),$$

*where  $M_2$  is the standard Hardy-Littlewood maximal operator along vertical line segments.*

**Claim 11.** *For  $\lambda > 0$ ,*

$$|\{x \in A_k : M_2 T_{j,J,n}^*f(x) > \lambda\}| \lesssim 2^{-|j-k|} |\{x \in \mathbb{R}^2 : M_2 T_{j,J,n}^*f(x) > \lambda\}|.$$

With these two claims, we see that

$$\begin{aligned} \|T_k T_{j,J,n}^* f\|_2^2 &\lesssim \int_0^\infty \lambda 2^{-|j-k|} |\{x \in \mathbb{R}^2 : M_2 T_{j,J,n}^* f(x) > \lambda\}| d\lambda \\ &= 2^{-|j-k|} \|M_2 T_{j,J,n}^* f\| \lesssim 2^{-|j-k|} \|f\|_2^2, \end{aligned}$$

since  $M_2$  and  $T_{j,J,n}$  are bounded on  $L^2$  with uniform constants, which proves estimate (8), and hence the estimate (3). ( $T_{j,J,n}$  is bounded because each  $\mathcal{R}_n^j$  is a good collection.) We turn to the proofs of these two claims.

*Proof of Claim 10.* Fix any  $K \in \mathcal{I}_k$  such that  $K \subseteq J$ . (If  $x \notin J \times [0, 1]$ , then  $T_k T_{j,J,s}^* f(x) = 0$ .) There exists  $J' \in \mathcal{I}_{j+1}$  with  $K \subseteq J' \subseteq J$ . Note that all  $R \in \mathcal{R}_n^j$  with  $\pi_1(R) \supseteq J'$ , have the same slope. For suppose such  $R_1, R_2$  have different slopes. Then by the fact mentioned at the beginning of the proof of Claim 9, and by the definition of the sets  $F_n$  given in the last section, we know  $\pi_1(R_1) \cap \pi_1(R_2) = \emptyset$ . But this contradicts the claim that  $\pi_1(R_1)$  and  $\pi_1(R_2)$  both contain  $J'$ .

Hence all  $R \in \mathcal{R}_n^j$  with  $\pi_1(R) \supseteq J'$ , have the same slope; let's call it  $\theta$ . This implies that  $T_{j,J,n}^* f(x)$  is constant along line segments contained in  $J'$  with slope  $\theta$ . Let  $Y_\theta$  be any line orthogonal to a line segment with slope  $\theta$ . Because  $T_{j,J,n}^* f(x)$  is constant along line segments contained in  $J' \times [0, 1]$  with slope  $\theta$ , we know that if  $L$  is a line segment contained in  $J' \times [0, 1]$ , then

$$\frac{1}{|L|} \int_L T_{j,J,n}^* f(x)$$

depends only on the projection of  $L$  onto the axis  $Y_\theta$ , and in particular, it does not depend on the slope of  $L$ . (Of course the integral here is with respect to one-dimensional Lebesgue measure.) Hence

$$\frac{1}{|L|} \int_L T_{j,J,n}^* f(x) \leq M_2 T_{j,J,n}^* f(x).$$

Since  $T_k$  is essentially an average over line segments, this finishes the proof of the claim.  $\square$

*Proof of Claim 11.* The set  $A_k \cap (J \times [0, 1])$  is supported on the set

$$\bigcup_{J' \in \mathcal{I}_{j+1}} (J' \times [0, 1]).$$

Hence it suffices to prove that for any  $J' \in \mathcal{I}_{j+1}$ , we have

$$\begin{aligned} |\{x \in A_k \cap (J' \times [0, 1]) : M_2 T_{j,J,n}^* f(x) > \lambda\}| &\lesssim \\ 2^{-|j-k|} |\{x \in (J' \times [0, 1]) : M_2 T_{j,J,n}^* f(x) > \lambda\}|. \end{aligned}$$

So we fix attention on a particular  $J' \in \mathcal{I}_{j+1}$ . By the argument in the proof of the previous claim, we know that all  $R \in \mathcal{R}_n^j$  with  $\pi_1(R) \supseteq J'$ , have the same

slope  $\theta$ , which implies that  $T_{j,J,n}^*f(x)$  is constant along line segments contained in  $J'$  with slope  $\theta$ . This further implies that  $M_2T_{j,J,n}^*f(x)$  is constant along segments of length  $|J'|$  with slope  $\theta$ . But since

$$|J' \cap \mathbf{sh}(\mathcal{I}_k)| \lesssim 2^{-|j-k|}|J'|,$$

we have proved the claim.  $\square$

## 6 Proof of Lemma 6

In this section, we prove Theorem 6. The argument given here is due to Katz [3]. Recall that we assume  $\mathcal{R}_1, \mathcal{R}_2, \dots, \mathcal{R}_N$  are good collections of rectangles and  $\mathcal{R} = \cup_{j=1}^N \mathcal{R}_j$ . To prove the theorem, we prove the weak-type estimate

$$|\{M_{\mathcal{R}}f > \lambda\}| \lesssim \log N \frac{\|f\|_2^2}{\lambda^2}.$$

To prove the weak-type bound above, we linearize the maximal operator as above, and prove restricted strong-type bounds for the linearization. That is, we prove

$$\|T^* \mathbf{1}_E\|_2^2 \lesssim \log N |E|$$

for any set  $E$ , where again we write  $T$  to denote a particular linearization of  $M_{\mathcal{R}}$ . As before, we will let  $\rho: [0, 1]^2 \rightarrow \mathcal{R}$  denote the linearization. Of course the estimates are independent of the particular linearization.

To upgrade this weak-type estimate to the desired strong-type estimate, one only needs to apply standard interpolation theorems. (Since  $T$  is trivially bounded on  $L^\infty$ , we may interpolate to obtain strong-type estimates for  $p > 2$ . Then interpolate with the trivial weak (1,1) estimate of  $\sim N$  to obtain the claimed strong bounds on  $L^2$ .)

For the rest of the section, we focus on proving this restricted strong-type estimate for  $T^*$ . It is convenient to assume that  $\pi_1(R)$  is a dyadic interval; we do so. For a set  $F$  and any interval  $I$ , let

$$F_I = \{x \in F: \pi_1(\rho(x)) \subseteq I\}.$$

For any rectangle  $R$  and any set  $F$ , define

$$B_R^F = \frac{1}{|R|} \int_R T^*(\mathbf{1}_{F_{\pi_1(R)}}).$$

and

$$\nu_R^F = |\{x \in F: \rho(x) = R\}|.$$

(Recall that  $\pi_1(R)$  is the projection of  $R$  onto the horizontal axis.) The quantity  $B_R$  is called the *badness* of the rectangle  $R$ . Before we proceed any further, we present one computation that is crucial for understanding this section.

**Claim 12.** *For a set  $F$ ,*

$$T^*(\mathbf{1}_F)(x) = \sum_{R \in \mathcal{R}} \frac{\nu_R^F \mathbf{1}_R(x)}{|R|}.$$

This is a weighted count of the rectangles in  $\mathcal{R}$  that contain  $x$ .

*Proof.* Recall that

$$Tf(x) = \int \frac{\mathbf{1}_{\rho(x)}(y)}{|\rho(x)|} f(y) dy.$$

This means that

$$\begin{aligned} T^*(\mathbf{1}_F)(x) &= \int \frac{\mathbf{1}_{\rho(y)}(x)}{|\rho(y)|} \mathbf{1}_F(y) dy \\ &= \sum_{R \in \mathcal{R}} \int_{\{y: \rho(y)=R\}} \frac{\mathbf{1}_{\rho(y)}(x)}{|\rho(y)|} \mathbf{1}_F(y) dy \\ &= \sum_{R \in \mathcal{R}} \frac{\mathbf{1}_R(x)}{|R|} \int_{\{y: \rho(y)=R\}} \mathbf{1}_F(y) dy \\ &= \sum_{R \in \mathcal{R}} \frac{\nu_R^F \mathbf{1}_R(x)}{|R|}. \end{aligned}$$

□

An immediate corollary of this is the estimate

$$\int T^*(\mathbf{1}_F) = \sum_{R \in \mathcal{R}} \nu_R^F \leq |F|. \quad (9)$$

Because of this computation, we see that the badness  $B_R$  is a weighted count of the rectangles  $R'$  that intersect  $R$  and that are essentially shorter than  $R$ . The weighting depends on the measure  $\nu_{R'}^F$ , the length of  $R'$ , and the angle between  $R$  and  $R'$ . (If  $R$  and  $R'$  intersect with smaller angle, then  $\mathbf{1}_{R'}(x)$  will be supported on a larger portion of  $R$ .)

**Claim 13.**

$$\int (T^*(\mathbf{1}_E))^2 \lesssim \sum_{R \in \mathcal{R}} \nu_R^E B_R.$$

Because of this claim and the definition of  $B_R$ , we see that  $\|T^*(\mathbf{1}_E)\|_2^2$  is essentially a count of quantities like  $|R_1 \cap R_2|$ , which is to be expected in an  $L^2$  estimate of this operator. This claim follows from a straightforward computation which we carry out shortly. We will combine it with the following lemma to prove the theorem.

**Lemma 14.** *Let  $\mathcal{R}_0$  be a good collection of rectangles. Let*

$$\mathcal{S}_k = \{R \in \mathcal{R}_0 : B_R \in [k-1, k)\}.$$

*Then*

$$|\bigcup_{R \in \mathcal{S}_k} R| \lesssim 2^{-ck}|E|.$$

We can already use these facts to prove the theorem of this section. Let

$$\mathcal{S}_{n,k} = \{R \in \mathcal{R}_n : B_R \in [k-1, k)\}.$$

By the claim,

$$\begin{aligned} \int (T^*(\mathbf{1}_E))^2 &\lesssim \sum_{R \in \mathcal{R}} \nu_R^E B_R \\ &= \sum_{n=1}^N \sum_{R \in \mathcal{R}_n} \nu_R^E B_R \\ &= \sum_{k=1}^{\infty} \sum_{n=1}^N \sum_{R \in \mathcal{S}_{n,k}} \nu_R^E B_R \\ &\lesssim \sum_{k=1}^{\infty} k \sum_{n=1}^N \sum_{R \in \mathcal{S}_{n,k}} \nu_R^E \\ &= (\star) \end{aligned}$$

Note that  $\sum_{R \in \mathcal{S}_{n,k}} \nu_R^E \leq |\bigcup \mathcal{S}_{n,k}|$  and that  $\sum_{R \in \mathcal{R}} \nu_R^E \leq |E|$ . This first estimate is useful when  $k$  is large, and the second when  $k$  is small. So by the lemma,

$$\begin{aligned} (\star) &\lesssim \sum_{k=1}^{\sim \log N} \log N \sum_{n=1}^N \sum_{R \in \mathcal{S}_{n,k}} \nu_R^E + \sum_{k \sim \log N}^{\infty} k \sum_{n=1}^N \sum_{R \in \mathcal{S}_{n,k}} \nu_R^E \\ &\lesssim \log N |E| + \sum_{k \sim \log N}^{\infty} k N 2^{-ck} |E| \\ &\lesssim \log N |E|. \end{aligned}$$

This proves Theorem 6 modulo Claim 13 and Lemma 14. First we prove Claim 13.

*Proof of Claim 13.* The proof is a straightforward computation:

$$\begin{aligned}
\int (T^*(\mathbf{1}_E))^2 &= \int \left( \sum_{R \in \mathcal{R}} \frac{\nu_R^F \mathbf{1}_R(x)}{|R|} \right)^2 \\
&= \int \sum_{R \in \mathcal{R}} \sum_{Q \in \mathcal{R}} \frac{\nu_R^F \mathbf{1}_R(x)}{|R|} \frac{\nu_Q^F \mathbf{1}_Q(x)}{|Q|} \\
&\lesssim \sum_{R \in \mathcal{R}} \nu_R^F \frac{1}{|R|} \int_R \sum_{Q \in \mathcal{R}: \pi_1(Q) \subseteq \pi_1(R)} \frac{\nu_Q^F \mathbf{1}_Q(x)}{|Q|},
\end{aligned}$$

where we have used symmetry to restrict the sum in the final integral to rectangles  $Q$  that are essentially shorter than  $R$ . To finish the proof, we need only note that

$$\sum_{Q \in \mathcal{R}: \pi_1(Q) \subseteq \pi_1(R)} \frac{\nu_Q^F \mathbf{1}_Q(x)}{|Q|} = T^*(\mathbf{1}_{F_{\pi_1(R)}}).$$

□

To prove Lemma 14, we iterate the following key lemma.

**Lemma 15.** *Let  $E$  be a set. Let  $\mathcal{R}_0$  be a good collection of rectangles. There exists a set  $E'$  such that  $|E'| \leq \frac{1}{2}|E|$  and such that for all  $R \in \mathcal{R}$ , either*

$$B_R^E \leq C$$

or

$$R \subseteq E' \text{ and } B_R^E \leq C + B_R^{E'}.$$

Here  $C$  is a universal constant.

## 6.1 Proof that Lemma 15 implies Lemma 14

Given Lemma 15, we proceed as follow. Define  $E_0 = E$ . Apply the lemma to find  $E_1$  with  $|E_1| \leq \frac{1}{2}|E_0|$  such that for every  $R \in \mathcal{R}_0$ , either

$$B_R^E \leq C,$$

or  $R \subseteq E'$  and

$$B_R^{E_0} \leq B_R^{E_1} + C.$$

Repeat to find  $E_2, E_3, E_4, \dots$ , with  $|E_{j+1}| \leq \frac{1}{2}|E_j|$  such that for every  $R \in \mathcal{R}_0$ , either

$$B_R^{E_j} \leq C$$

or  $R \subseteq E_{j+1}$  and

$$B_R^{E_j} \leq B_R^{E_{j+1}} + C.$$

Now suppose  $R$  is such that

$$B_R^{E_0} \geq Ck$$

for some integer  $k \geq 1$ . Then we know  $R \subseteq E_1$  and

$$B_R^{E_0} \leq C + B_R^{E_1}.$$

This implies

$$B_R^{E_1} \geq C(k-1).$$

Similarly, this implies  $R \subseteq E_2$ ,  $B_R^{E_0} \leq 2C + B_R^{E_2}$ , and

$$B_R^{E_2} \geq C(k-2).$$

Iterating, we see that  $R \subseteq E_{k-1}$ . This implies

$$|\bigcup_{R \in \mathcal{S}_{Ck}} R| \lesssim 2^{-k}|E|,$$

which proves Lemma 14.

## 6.2 Proof of Lemma 15

Fix any intervals  $I, K$ . Without loss of generality, we assume that the rectangles in  $\mathcal{R}_0$  that project vertically to  $I$  have slope zero. This is a notational convenience only. We need some notation to define the set  $E'$  from the statement of the lemma. In the following definitions, dependence on the set  $E$  is suppressed. Define

$$\mathcal{A}_{I,K}^{\text{in}} = \{R: \pi_1(R) \subseteq I \text{ and } \pi_2(R) \subseteq 3K\},$$

$$\mathcal{A}_{I,K}^{\text{out}} = \{R: \pi_1(R) \subseteq I \text{ and } \pi_2(R) \not\subseteq 3K\},$$

$$E_{I,K}^{\text{in}} = \{x \in E: \rho(x) \in \mathcal{A}_{I,K}^{\text{in}}\},$$

$$E_{I,K}^{\text{out}} = \{x \in E: \rho(x) \in \mathcal{A}_{I,K}^{\text{out}}\},$$

$$B_{I,K}^{\text{in}} = \frac{1}{|I||K|} \int_{I \times K} T^*(\mathbf{1}_{E_{I,K}^{\text{in}}}),$$

$$B_{I,K}^{\text{out}} = \frac{1}{|I||K|} \int_{I \times K} T^*(\mathbf{1}_{E_{I,K}^{\text{out}}}).$$

Note that for any axis parallel rectangle  $R$  with  $\pi_1(R) = I$ , we have

$$B_R^E = \frac{1}{|R|} \int_R T^*(\mathbf{1}_{E_{I,K}^{\text{in}}}) + \frac{1}{|R|} \int_R T^*(\mathbf{1}_{E_{I,K}^{\text{out}}})$$

for any interval  $K$ .

Let  $\mathcal{B}_I$  be the collection of intervals  $K$  such that

$$B_{I,K}^{\text{out}} \geq \lambda_0,$$

but such that

$$B_{I,3K}^{\text{out}} < \lambda_0$$

where  $\lambda_0 \geq 1$  is a universal constant to be specified later. We now define the set  $E'$  from the statement of Lemma 15: let

$$E' = \bigcup_I \bigcup_{K \in \mathcal{B}_I} (I \times 3K).$$

We also define the auxiliary set

$$F = \{x: M_{\mathcal{R}_0} T^* \mathbf{1}_E \geq \frac{\lambda_0}{2}\}.$$

We will show that  $|E'| \leq C|F|$  and then that  $|F| \leq \frac{|E|}{2C}$ . To prove the second estimate we need only the weak (1,1) estimate for  $M_{\mathcal{R}_0}$ . To prove the first estimate we need the following claim.

**Claim 16.** *If  $K \in \mathcal{B}_I$ , and  $\pi_1(R) = I$  and  $\pi_2(R) \subseteq K$ , then*

$$\frac{1}{|R|} \int_R T^*(\mathbf{1}_{E_{I,K}^{\text{out}}}) \leq 20\lambda_0.$$

*Proof.* We will show that for any  $a \in K$ ,

$$\frac{1}{|I|} \int_{I \times a} T^*(\mathbf{1}_{E_{I,K}^{\text{out}}}) \leq 20\lambda_0,$$

which implies the claim. (Of course the integration immediately above is with respect to 1-D Lebesgue measure.) Note that if  $R \in \mathcal{A}_{I,K}^{\text{out}}$  and  $R$  intersects  $I \times a$ , then  $R$  intersects  $I \times a'$  for  $a'$  in a set of measure  $\frac{1}{3}|3K|$ . This implies that if  $R \in \mathcal{A}_{I,K}^{\text{out}}$ , then

$$\frac{1}{|I||3K|} \int_{I \times 3K} \mathbf{1}_R \geq \frac{1}{10} \frac{1}{|I|} \int_{I \times a} \mathbf{1}_R$$

for every  $a \in K$ . Also note that

$$\int T^*(\mathbf{1}_{E_{I,K}^{out} \setminus E_{I,3K}^{out}}) \leq \int T^*(\mathbf{1}_{I \times 3K}) \leq |I||3K|$$

by (9), because  $T^*$  is positive and because  $E_{I,K}^{out} \setminus E_{I,3K}^{out} \subseteq I \times 3K$ . Combining this with the fact that  $B_{I,3K}^{out} < \lambda$ , we know

$$\begin{aligned} \frac{1}{|I|} \int_{I \times a} T^*(\mathbf{1}_{E_{I,K}^{out}}) &\leq 10 \frac{1}{|I||3K|} \int_{I \times 3K} T^*(\mathbf{1}_{E_{I,K}^{out}}) \\ &= 10 \frac{1}{|I||3K|} \int_{I \times 3K} T^*(\mathbf{1}_{E_{I,K}^{out} \setminus E_{I,3K}^{out}}) \\ &+ 10 \frac{1}{|I||3K|} \int_{I \times 3K} T^*(\mathbf{1}_{E_{I,3K}^{out}}) \\ &\leq 10 + 10 \frac{1}{|I||3K|} \int_{I \times 3K} T^*(\mathbf{1}_{E_{I,3K}^{out}}) \\ &\leq 20\lambda_0. \end{aligned}$$

□

Consider an interval  $I$  with  $K \in \mathcal{B}_I$ . By the proof of the previous claim, we know that there exists  $A \subseteq K$ , such that  $|A| \geq \frac{1}{20}|K|$ , and such that

$$\frac{1}{|I|} \int_{I \times a} T^*(\mathbf{1}_{E_{I,K}^{out}}) \geq \frac{\lambda_0}{2}$$

for all  $a \in A$ , where here the integral is taken with respect to one-dimensional measure on  $I \times a$ . This is because the proof gives an upper bound on such averages; this, together with the lower bound on  $B_{I,K}^{out}$  yields the claimed lower bound for many  $a \in A$ . Hence if  $a \in A$ , then for all  $x \in I \times a$ , we have

$$M_{\mathcal{R}_0} T^*(\mathbf{1}_E)(x) \geq \frac{\lambda_0}{2}.$$

This implies that if  $x \in (I \times 3K)$ , then

$$M_2 \mathbf{1}_F(x) \geq \frac{1}{40}.$$

Hence

$$\begin{aligned} |E'| &= \left| \bigcup_I \bigcup_{K \in \mathcal{B}_I} (I \times 3K) \right| \\ &\leq |\{x: M_2 \mathbf{1}_F(x) \geq \frac{1}{40}\}| \\ &\leq C|F| \\ &= C|\{x: M_{\mathcal{R}_0} T^*(\mathbf{1}_E) \geq \frac{\lambda_0}{2}\}| \\ &\leq \frac{C}{\lambda_0} \|(T^* \mathbf{1}_E)\|_1 \leq \frac{1}{2}|E|, \end{aligned}$$

provided  $\lambda_0$  is large enough. Here we used the fact that  $\mathcal{R}_0$  is a good collection of rectangles.

We just proved the claim about the size of  $E'$ . If  $B_R \geq \lambda_0$ , then there exists  $K$  such that  $K \in \mathcal{B}_{\pi_1(R)}$ . Further, by Claim 16 and the positivity of  $T^*$ ,

$$\begin{aligned} B_R &= \frac{1}{|R|} \int_R T^* \mathbf{1}_{E_{I,K}^{\text{in}}} + \frac{1}{|R|} \int_R T^* \mathbf{1}_{E_{I,K}^{\text{out}}} \\ &\leq B_R^{E'} + 20\lambda_0, \end{aligned}$$

which proves the other claim of Lemma 15.

## References

- [1] Bateman, Michael.  *$L^p$  estimates for maximal averages along one-variable vector fields in  $\mathbf{R}^2$* , Proc. Amer. Math. Soc. **137**, (2009), 955-963.
- [2] Duoandikoetxea, J. *Fourier Analysis*, 2001, AMS Graduate Studies in Mathematics, vol. 29
- [3] Katz, N.H. *Maximal operators over arbitrary sets of directions*, Duke Math. J. vol. 97, no. 1 (1999), 67-79.
- [4] Katz, N.H. *Remarks on maximal operators over arbitrary sets of directions*, Bull. London Math. Soc. vol. 31, 1999, no. 6, pages 700-710
- [5] Lacey, Michael, and Xiaochun Li. *Maximal Theorems for the Directional Hilbert Transform on the Plane* Trans. Amer. Math. Soc. 358 (2006), 4099-4117.
- [6] Lacey, Michael, and Xiaochun Li. *On a Conjecture of EM Stein on the Hilbert Transform on Vector Fields* Memoirs of the AMS 205 (2010), no. 965.
- [7] Lacey, Michael, and Xiaochun Li. *On a Lipschitz Variant of the Kakeya Maximal Function* Available at <http://arxiv.org/abs/math/0601213>